Journal of Hydrology 394 (2010) 275–284

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Towards probabilistic forecasting of flash floods: The combined effects of uncertainty in radar-rainfall and flash flood guidance

Gabriele Villarini [a,*], Witold F. Krajewski [b], Alexandros A. Ntelekos [a], Konstantine P. Georgakakos [c], James A. Smith [a]

[a] Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ 08540, USA
[b] IIHR-Hydroscience & Engineering, The University of Iowa, Iowa City, IA, USA
[c] Hydrologic Research Center, San Diego, CA, USA

ARTICLE INFO

Keywords:
Flash flood
Flash flood guidance
Radar-rainfall
Uncertainties

SUMMARY

The flash flood guidance system (FFGS) is an operational system which assists forecasters to issue flash flood warnings and watches over the conterminous United States. Currently, it is a deterministic system: given a specified precipitation accumulation over a basin and over a time period, issuing of flash flood watches and warnings is considered depending on the exceedance of a certain threshold value (flash flood guidance – FFG). For given channel characteristics and initial soil moisture conditions, FFG values are computed with the use of a hydrologic model.

The purpose of this study is to consider the effects of radar-rainfall and flash flood guidance uncertainties on the FFGS. The errors in the FFG are accounted for by quantifying the uncertainties due to the estimation of the hydraulic and terrain characteristics, and the hydrologic model parameters and initial state. To account for the uncertainties in radar-rainfall, the authors use an empirically-based radar-rainfall error model. This model requires calibration for each application region and it accounts for range effects, synoptic conditions, space–time resolutions, and the spatial and temporal dependences of the errors. Thus, the combined effects of uncertainty in both radar-rainfall and FFG can be assessed. The results are exemplified through two cases in a small basin in the Illinois River Basin. The potential of transforming FFGS into a probabilistic system is also discussed.

Introduction

According to the American Meteorological Society (AMS; AMS, 2000), flash floods are characterized by small spatial and temporal scales, generally less than 1000 km^2 and less than 6 h. Despite the small space–time scales, flash floods "continue to be one of nature's worst killers" (AMS, 2000), taking a high human and economic toll (e.g., Senesi et al., 1996; Baeck and Smith, 1998; Chang, 1998; Petersen et al., 1999; Pontrelli et al., 1999; Ogden et al., 2000; Pielke and Downton, 2000; Smith et al., 2000, 2001; Weaver et al., 2000; Scofield and Kuligowski, 2003; O'Connor and Costa, 2004; Delrieu et al., 2005; Downton et al., 2005; NRC, 2005; Chancibault et al., 2006; Borga et al., 2007; Choi et al., 2007; Carpenter et al., 2007; Ashley and Ashley, 2008; Jessup and DeGaetano, 2008).

The study of forecast uncertainty for flash floods has gained momentum in the last few years (e.g., Krzysztofowicz, 2001; Ferraris et al., 2002; Siccardi et al., 2005; Taramasso et al., 2005; Georgakakos, 2006; Ntelekos et al., 2006; Collier, 2007; Reggiani and Weerts, 2008; Norbiato et al., 2008, 2009; Morin et al., 2009), since a deterministic system provides the user or the decision maker with an "illusion of certainty" (Krzysztofowicz, 2001). As acknowledged in the policy statement by AMS (2000), important research challenges include "quantifying forecast uncertainty by providing probabilistic forecast guidance" (AMS, 2000). However, the research in this field is still in its early stages.

This study presents a probabilistic framework that is applicable to the real-time flash-flood forecasting problem. In particular, we focus on the flash flood guidance system (FFGS) used by the National Weather Service (NWS) to assist with decisions regarding flash flood warnings and watches over the conterminous United States. In the current deterministic version of the FFGS system, the amount of rainfall of a given duration necessary to cause flooding conditions at the catchment outlet is compared with the corresponding radar-based rainfall estimates produced by the Next Generation Radar (NEXRAD) stations (e.g., Crum and Alberty, 1993; Klazura and Imy, 1993; Fulton et al., 1998). In brief, the system consists of two components: the main component is related to the estimation of the flooding conditions using hydrologic and hydraulic inputs for given initial soil moisture conditions. The second component is related to the operational estimation of radar-rainfall (see Ntelekos et al. (2006) for a more detailed

* Corresponding author.
E-mail address: gvillari@princeton.edu (G. Villarini).

0022-1694/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2010.02.014

description of the FFGS). The FFGS helps to decide whether a flash flood watch or warning should be issued by comparing the basin-average radar-rainfall estimate with respect to the flash flood guidance (*FFG*), which represents the rainfall necessary to cause flash flooding for given soil moisture conditions. As mentioned in Ntelekos et al. (2006), no action is taken if *FFG* is much larger than the observed or forecasted rainfall, while issuing a flash flood warning is considered if *FFG* is smaller than the rainfall.

Each component is affected by various sources of uncertainty that could/should be accounted for. Using different representations of the operational hydrologic model, Ntelekos et al. (2006) and Georgakakos (2006) characterized the combined uncertainties associated with the core part of the system, namely the calculation of the *FFG* values. These uncertainties included the combined effect of hydrologic, hydraulic and initial soil moisture conditions. Ntelekos et al. (2006) concluded that they can be effectively described by a lognormal distribution, for the cases examined. The authors also noted, "the problem of uncertainty quantification of the entire FFGS requires consideration of several additional components". In particular, they assumed that the radar-rainfall estimates were error-free, even though it is well-acknowledged that these estimates are affected by several sources of uncertainty (e.g., Wilson and Brandes, 1979; Krajewski and Smith, 2002; Germann et al., 2006; Ciach et al., 2007; consult Villarini and Krajewski, (2010a) for a recent literature review of this topic). To address this issue, Ciach et al. (2007) proposed an empirically-based error model able to account for uncertainties associated with radar-based estimates of rainfall. Based on this model, Villarini et al. (2009a) presented a generator of probable true rainfall fields conditioned on radar-rainfall maps. These results, in conjunction with the findings by Ntelekos et al. (2006), allow the transformation of the FFGS from deterministic to probabilistic, which accounts for the uncertainties due to radar-rainfall estimation, as well as the hydrologic and hydraulic components. This probabilistic framework would enable decision makers to make more informed decisions.

It is worth clarifying that in this study we are focusing on the real-time flash flood problem, in which the observed radar-rainfall is compared against *FFG* (see also Martina et al. (2006), Ntelekos et al. (2006), Norbiato et al. (2008), and Norbiato et al. (2009)). In forecasting of flash floods, quantitative precipitation estimates could be obtained by means of Numerical Weather Predictions models, short-term forecasting (or nowcasting) techniques, or using radar-rainfall as a persistence surrogate for future rainfall (e.g., Collier, 2007). The uncertainties in the selected forecasting technique would add to the uncertainties in the radar-rainfall estimates, making the flash-flood forecasting problem even more uncertain. Moreover, even though in operational conditions radar-rainfall estimates are used in the calculation of the initial soil moisture conditions, in this study, we do not investigate how radar-rainfall errors affect the soil moisture conditions.

The paper is organized in the following way. In the next section, we describe the models of the uncertainties in radar-rainfall estimates and *FFG*. In 'Data', we present two illustrative cases. In 'Results' these cases are analyzed under four different scenarios: (i) deterministic in *FFG* and radar-rainfall (current); (ii) probabilistic in *FFG* and deterministic in radar-rainfall; (iii) deterministic in *FFG* and probabilistic in radar-rainfall; and (iv) probabilistic in both *FFG* and radar-rainfall. 'Discussion and conclusions' summarizes the main points and concludes the paper.

Modeling of uncertainties

In this section, we briefly describe the error models used to account for the uncertainties in both the *FFG* and the radar-rainfall estimates. For more details, the reader is pointed to Ntelekos et al. (2006), Ciach et al. (2007), and Villarini et al. (2009a).

Uncertainties in flash flood guidance

The NWS established the FFGS as a tool to assist forecasters with decisions pertaining to flash flood watches and warnings over the conterminous United States. Its backbone is represented by the threshold-runoff (*Thresh-R*) component, which corresponds to the amount of effective rainfall that is uniformly distributed over a certain basin and is capable of causing flooding conditions. Carpenter et al. (1999) computed it by equating the peak runoff rate obtained using the Geomorphologic Unit Hydrograph theory (e.g., Rodriguez-Iturbe and Valdes, 1979; Rodriguez-Iturbe et al., 1982a,b) to the bankfull flow based on Manning's steady uniform resistance formula. For a given rainfall duration (hourly, three-hourly, and six-hourly) *Thresh-R* is a one-time task and has been computed off-line for all of the basins for which FFGS is effective. By using the Sacramento Soil Moisture Accounting Model (SAC-SMA; Burnash, 1995) and initial soil moisture conditions provided by the River Forecasting Centers (RFCs) every 6 h, rainfall–runoff curves are generated in a "what if" scenario. The (future) rainfall that causes runoff equal to *Thresh-R* under specific soil moisture conditions is the *FFG* value of the basin, for a certain rainfall duration (e.g., Georgakakos, 2006; Ntelekos et al., 2006).

Ntelekos et al. (2006) focused on the characterization of the uncertainties associated with *FFG*. First, they showed that *Thresh-R* is a nonlinear function of two classes of data: those that can be derived by Geographical Information Systems (GIS), including drainage area, main stream length, and Horton's length ratio, and those derived from regional relationships (Carpenter et al., 1999), including the channel top width and the hydraulic depth at the outlet, and the local channel slope. To study the uncertainty associated with the *Thresh-R* calculation, uncertainty information should be obtained for those variables. For the first class of variables, there is very limited information on the shape of the error distribution. Consequently, Ntelekos et al. (2006) assumed generalized beta error distributions and decided the uncertainty bounds based on data from previous studies (Carpenter and Georgakakos, 1993). For the class of variables obtained from regional relationships, uncertainty information existed for the case of Oklahoma, and the error distribution of these variables (lognormal) was directly derived.

The total uncertainty in the *Thresh-R* calculation was propagated in the FFGS with the use of a simplified version of the SAC-SMA model. Additional uncertainty in the model parameters and initial state was assumed to account for total uncertainty of the simplified FFGS. The authors showed that the total uncertainty in the *FFG* calculation could be well described by a lognormal distribution under two initial soil moisture conditions (dry and wet).

Uncertainties in radar-rainfall estimates

Ciach et al. (2007) proposed an empirically-based radar-rainfall error model, in which the relation between true ground areal rainfall R_a and radar-rainfall R_r was described by the product of a systematic distortion function and random component:

$$R_a = h(R_r)\varepsilon(R_r) \tag{1}$$

where the systematic distortion function $h(R_r)$ accounts for systematic errors (bias function of the radar-rainfall estimates), while the random component $\varepsilon(R_r)$ describes the remaining random uncertainties. Both components of the model are a function of radar-rainfall values, and the true rainfall was approximated by rain gauge measurements.

As a preliminary step, the overall bias (defined as the ratio between rain gauge and radar sample means) was computed and removed (it depends on the distance from the radar and season, used as proxy for rainfall regime). Using a non-parametric approach, the

systematic component was computed as an expectation function conditioned on the radar-rainfall estimates (see also Villarini et al. (2008)). Once h has been computed, the calculation of the random component is straightforward. To obtain a more concise description of both the systematic and random components, Ciach et al. (2007) approximated the radar-rainfall uncertainty model with parametric models. They found that the systematic distortion function can be parameterized by a power law function. They also found that the random component could be described by a Gaussian distribution with mean equal to 1, a standard deviation that is a monotonically-decreasing (hyperbolic) function of radar-rainfall values, and with spatio-temporal dependencies that can be well described by a two-parameter exponential correlation function (Villarini et al., 2009a). Therefore, all of this error model's components have simple parametric representations. To account for range effects, Ciach et al. (2007) divided the radar umbrella into five zones, while they accounted for synoptic conditions by stratifying the sample into three seasons: cold (January, February, March, November, and December), warm (April, May, and October), and hot (June, July, August, and September). The results presented by Ciach et al. (2007) are based on a large sample (6 years) of radar data from the NEXRAD site in Oklahoma City (KTLX), hourly and larger time scales, averaged over the Hydrologic Rainfall Analysis Project (HRAP) grids (approximately 4 km by 4 km pixels; Reed and Maidment, 1999), and NEXRAD reflectivity–rainfall (Z–R) relation (Fulton et al., 1998). The radar estimates are complemented with rainfall measurements by the Oklahoma Micronet (e.g., Allen and Naney, 1991) and Mesonet (e.g., Brock et al., 1995) rain gauge networks.

Building on the results by Ciach et al. (2007), Villarini et al. (2009a) developed a generator of ensembles of probable true rainfall fields conditioned on a given radar-rainfall map. Taking advantage of the Gaussian nature of the random component, it is also possible to create probabilistic maps, which provide information on the probability that the true rainfall is larger than a certain threshold of interest R_{thres}. Since we are interested in the probability of exceedance of the flash flood guidance by the true rainfall over a certain basin, we use the following analytical expression from Villarini et al. (2009a):

$$P\left(\frac{1}{n}\sum_{i=1}^{n} R_{a,i} \geqslant R_{thres}\right) = 1-\Phi(x) = 1-\Phi\left(\frac{R_{thres}-\frac{\sum_{i=1}^{n} h_i}{n}}{\frac{1}{n}\sqrt{\sum_{i=1}^{n} h_i^2\sigma_{\varepsilon,i}^2+\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\rho_{s,ij}h_i h_j\sigma_{\varepsilon,i}\sigma_{\varepsilon,j}}}\right) \tag{2}$$

where i and j are indexes referring to certain pixels among the n number of pixels within the area of interest (flash flood prone catchment in this case), R_{thres} is the flash flood guidance value for a certain basin and soil moisture conditions, σ_ε and ρ_s are, respectively, the standard deviation and the spatial correlation of the random error component, and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

This error model has already been applied towards improving statistical validation of satellite precipitation estimates (Villarini et al., 2009b) and to investigate the impact of radar-rainfall uncertainties on the scaling properties of rainfall (Villarini and Krajewski, 2009b).

Combining radar-rainfall and FFG uncertainties

As mentioned before, in the current deterministic FFGS the *FFG* value is compared to the basin-average radar-rainfall estimates to help decide whether (or not) to issue a flash flood watch or warning. When we consider only the uncertainties in *FFG* and a deterministic radar-rainfall value, we can compute the probability that basin-average rainfall is larger than the *FFG* value. The same can be done if we account for the uncertainties in radar-rainfall and consider the *FFG* value as deterministic. However, when accounting for uncertainties in both of the components of the FFGS, we compare quantiles q_{FFG} from the *FFG* distribution against quantiles q_{RR} from the radar-rainfall distribution. For any pair (q_{FFG}, q_{RR}), we can check whether the *FFG* value is larger than the radar-rainfall value. If decisions are based only on the comparison between *FFG* and radar-rainfall values, we can create warning-no warning plots depending on whether the *FFG* value is larger than the radar-rainfall one.

However, users and forecasters will need to make deterministic decisions regarding whether to issue a flash flood warning or watch. The selection of an *FFG* and radar-rainfall quantile is dictated by both economic and societal issues. The selection of a high

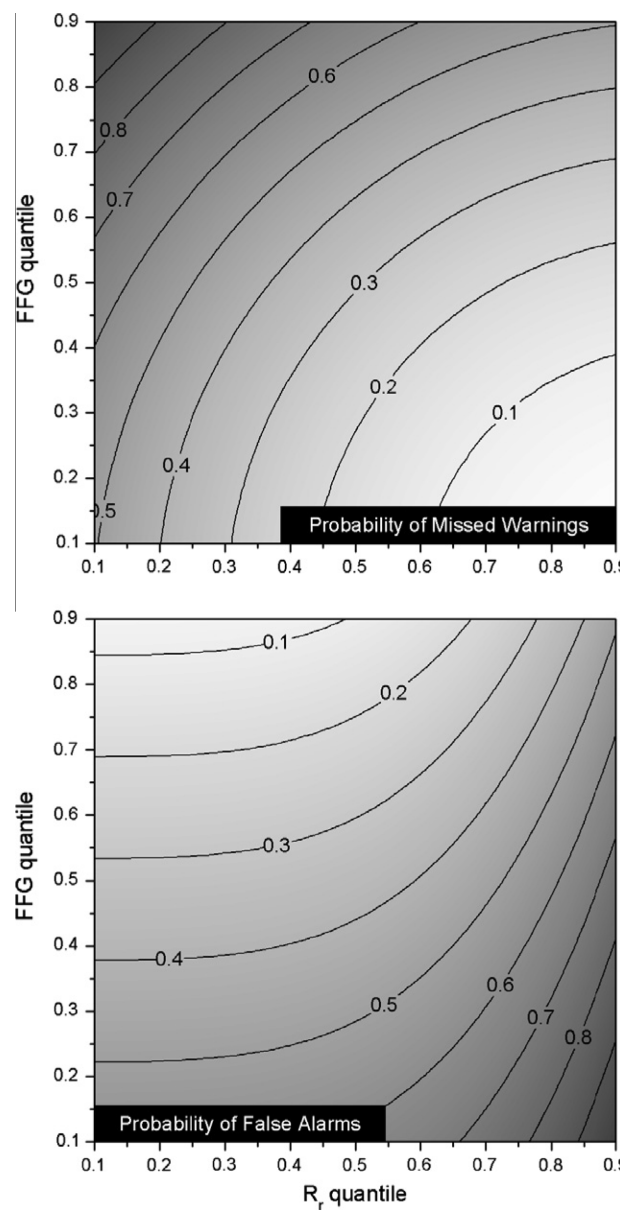

Fig. 1. Plots with the probability of missed warnings (upper panel) and the probability of false alarms (bottom panel) for different *FFG* and radar-rainfall quantiles.

FFG quantile leads to a higher probability of missed warnings (a flash flood occurred but no flash flood warning or watch was issued). This situation would lead to a high toll in terms of fatalities and economic damage. On the other hand, a high threshold in radar-rainfall quantiles would result in an increasing probability of false alarms, with high economical and societal repercussions. A possible way of providing suggestions concerning the selection of q_{FFG} and q_{RR} is by looking at the probability of false alarms (PFA) and the probability of missed warnings (PMW). Given the total number of warnings, the former can be defined as the probability of a warning if no event occurred. On the other hand, the probability of missed warnings can be defined as ratio between the number of missed warnings and the total number of flash floods. We would expect a higher probability of false alarms for increasing q_{RR} values, and a higher probability of missed warnings for increasing q_{FFG}.

For each basin and combination (q_{FFG}, q_{RR}) it would possible to compute these two probabilities from the data. In this study, we do not have this information and therefore we have created plots with PFA and PMW for different quantiles and we have summarized the results in Fig. 1. In this case, the variation of PMW and PFA as a function of q_{FFG} and q_{RR} is arbitrary, even though it reflects the idea that we would have more false alarms for higher q_{RR} (and/or lower q_{FFG}), and more missed warnings for higher q_{FFG} (and/or lower q_{RR}).

Similar to Martina et al. (2006), we define the "cost" as the damage perception by the user or forecaster. This "cost" does not directly translate into the economic damage associated with a flash flood, but includes both economic as well as unquantifiable damage (e.g., societal perception of false alarms). It is the task of stakeholders and decision makers to quantify the "cost" associated with false alarms and missed warnings. In this study, we assume that this piece of information is available. Following the definition of risk (it is the product of the probability of occurrence of a certain event and the "cost" associated with that event), for any pair (q_{FFG}, q_{RR}) we can compute the risk associated with a false alarm and with a missed warning. We can then define the risk associated with a flash flood (*riskFF*) as the sum of the risk associated with a missed warning and a false alarm. Following Martina et al. (2006), we can then select the most "convenient" *FFG* and radar-rainfall quantiles by minimizing *riskFF*.

Table 1
Scenarios of the initial soil moisture conditions investigated (in % of the capacity).

	Soil moisture conditions	
	Dry	Wet
Upper zone tension water content (%)	70	90
Upper zone free water content (%)	60	85

Table 2
Summary of the values of the *FFG* (mm/3 h) and the basin-averaged radar-rainfall estimates (mm/3 h) for the two cases and the two soil moisture scenarios considered in this study.

	Case I		Case II	
	Dry	Wet	Dry	Wet
Flash flood guidance (mm/3 h)	80.87	57.64	80.87	57.64
Basin-averaged radar-rainfall (mm/3 h)	67.73	67.73	43.70	43.70

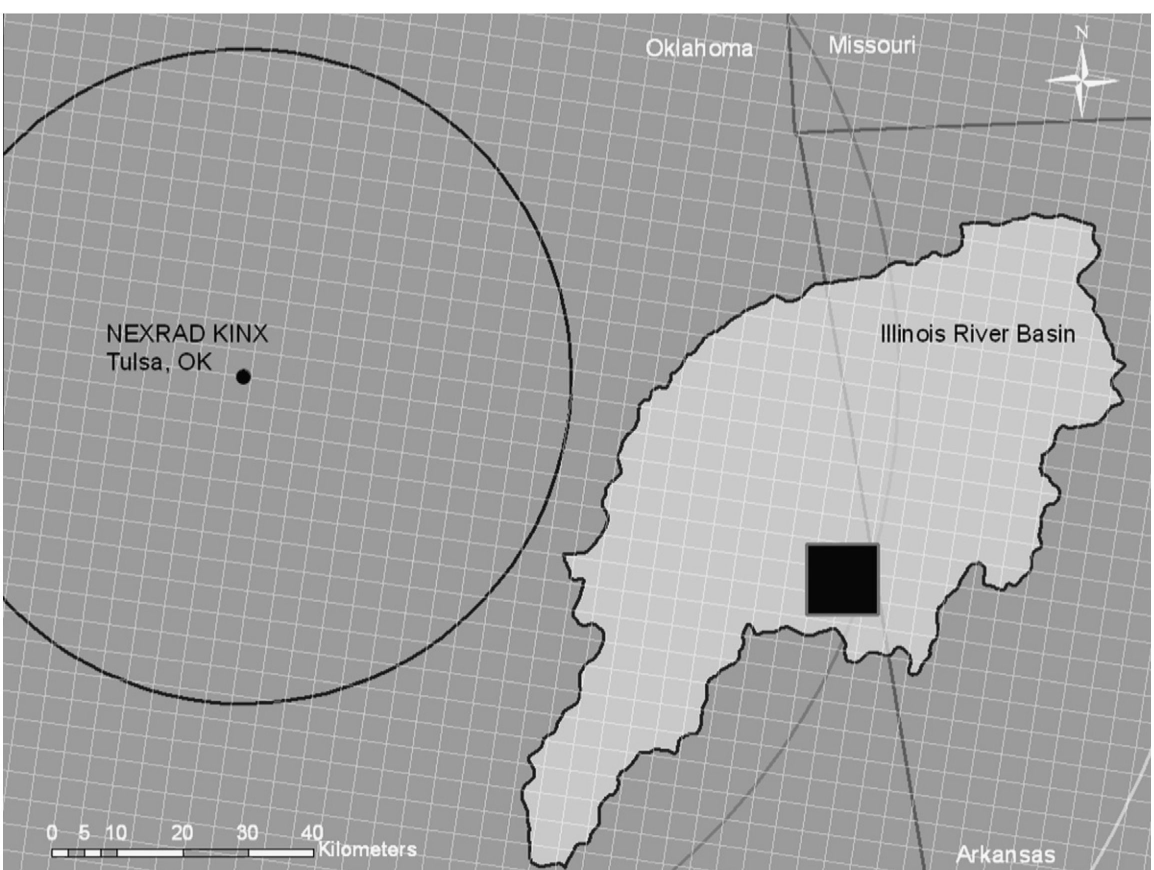

Fig. 2. Map of the Illinois River Basin and location of the basin under study (black box). The white grid refers to the HRAP grid. The radar rings centered on the Tulsa radar are 50 km apart.

As mentioned above, in this study, we assume as known the probabilities of missed warnings and false alarms for different combinations of *FFG* and radar-rainfall quantiles. In reality, these values could be obtained from the available record or by using a hydrologic model. Moreover, we also assume that the "cost" associated with false alarms and missed warnings was previously

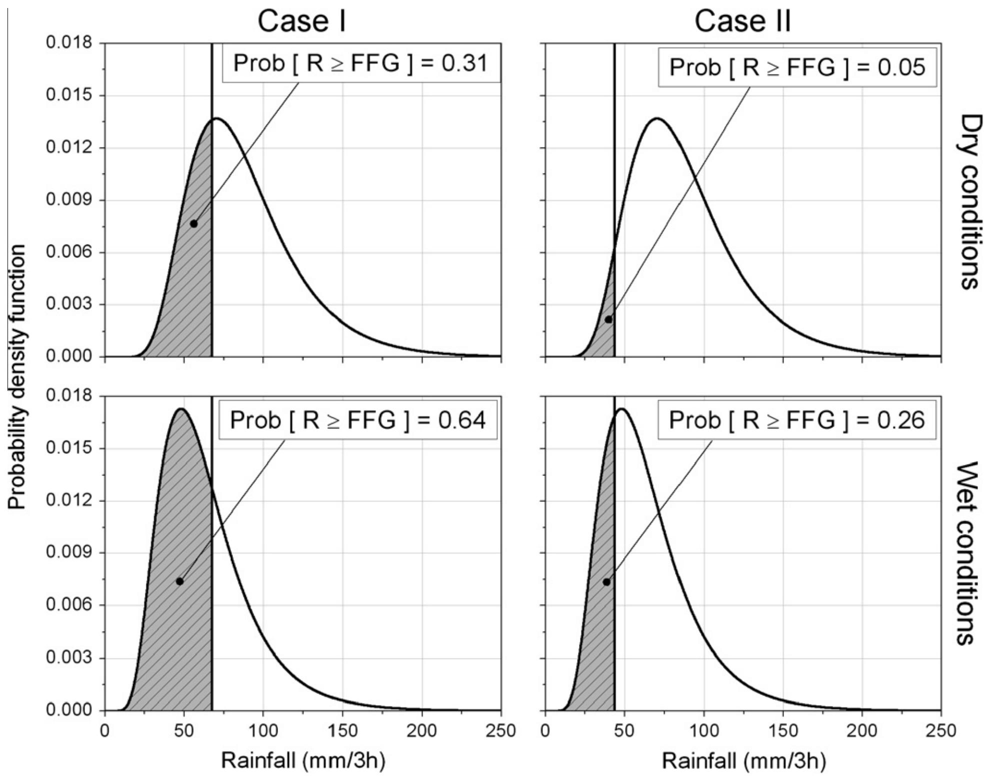

Fig. 3. Results from the application of a flash flood guidance system probabilistic in *FFG* and deterministic in R_r for the two cases under study and two different soil moisture conditions.

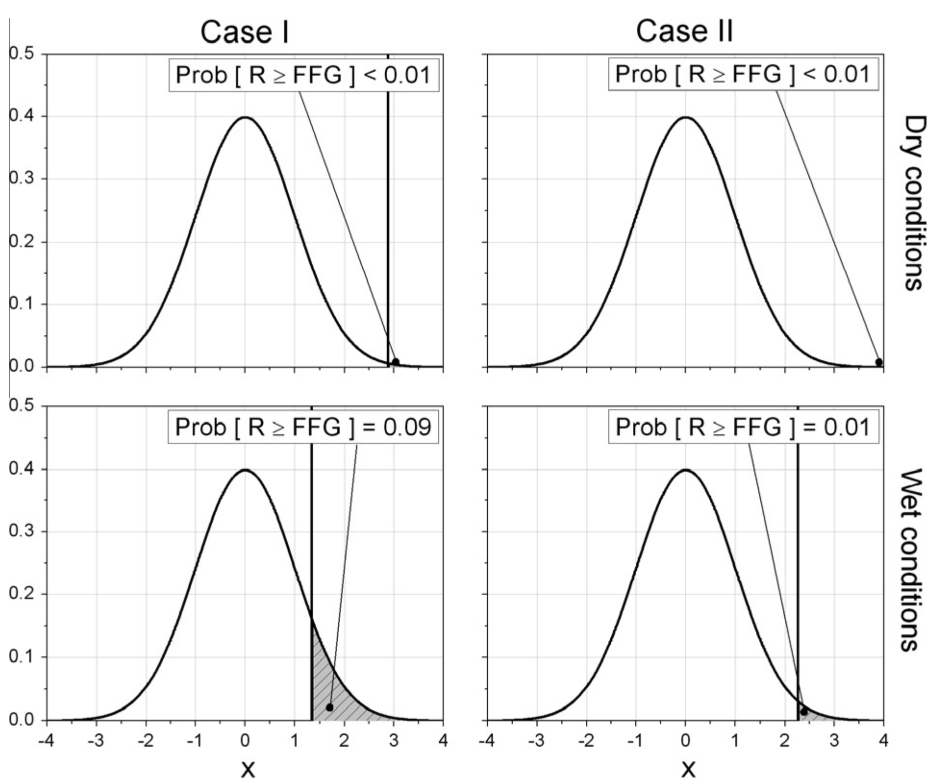

Fig. 4. Results from the application of a flash flood guidance system probabilistic in R_r and deterministic in *FFG* for the two cases under study and two different soil moisture conditions. On the abscissa, x refers to the quantity in Eq. (2).

quantified. For generality, we assume that the "cost" associated with false alarms is proportional to the cost associated with missed warnings. In this way, we can investigate how the most "convenient" *FFG* and radar-rainfall quantiles change as a function of "cost".

Data

In this study, we present the results of the methodology of the previous section for the same basin considered in Ntelekos et al. (2006). This catchment [US Geological Survey (USGS) basin cataloguing number 2568] is located about 100 km from the Tulsa (KINX) NEXRAD station (Fig. 2) and is a sub-basin of the Illinois River Basin, with a drainage area of about 116 km^2 and a stream length of 1.88 km. Since the selected basin is not covered by the Oklahoma City (KTLX) radar, we assume that the parameters of the radar-rainfall error model by Ciach et al. (2007) are valid for KINX radar as well. The Tulsa radar (an S-band radar like KTLX) is located at about 180 km from the Oklahoma City radar and orography is not a dominant factor in this entire area. Since the results in Ciach et al. (2007) are dependent on the selected radar-rainfall product and radar, there is no certainty that the parameters for the KTLX radar would be the same for the KINX radar, even though there is evidence that the overall model structure (in terms of systematic and random components) would be appropriate (see Villarini and Krajewski (2009a) for the application of this error model to a C-band radar in southern England). The transferability of the results in Ciach et al. (2007) to other radars, in particular KINX, should be the object of future works. As far as the soil moisture conditions are concerned, we investigate the same two scenarios (dry and wet conditions; Table 1) as in Ntelekos et al. (2006). Nonetheless, it is worth mentioning that different studies in the literature found that large uncertainties in the predictability of the hydrologic response were observed for intermediate soil moisture conditions when small soil moisture variations can produce significant changes on runoff (e.g., Zehe and Bloschl, 2004; Zehe et al., 2007; Zehe and Sivapalan, 2009).

We focus on the three-hourly scale and have selected two different storms: Case I which transpired on October 5th, 1998, and Case II which occurred on June 21st, 2000. The selection of these two cases is not associated with the occurrence of flash flooding over the selected catchment, but these two cases allow us to describe the modifications of the FFGS once uncertainties in both *FFG* and radar-rainfall are accounted for. The radar data was processed through the Hydro-NEXRAD software system (Krajewski, 2007a; Krajewski et al., 2008). It is a web-based system that allows generation of radar-rainfall products (derived from NEXRAD data from about 40 sites in the United States). Several customized modular radar-rainfall algorithms are available in Hydro-NEXRAD (Quick Look, High Res, Pseudo NWS PPS, and Custom; Krajewski, 2007b). In this study, we use the so-called "Pseudo-PPS," which is a suite of algorithms that tries to mimic the NWS Precipitation Processing System (PPS; Fulton et al., 1998) as closely as possible. As shown by Villarini and Krajewski (2010b), the radar-rainfall modeling results from "Pseudo-PPS" favorably compared with those by Ciach et al., 2007 using NWS PPS. We generated hourly

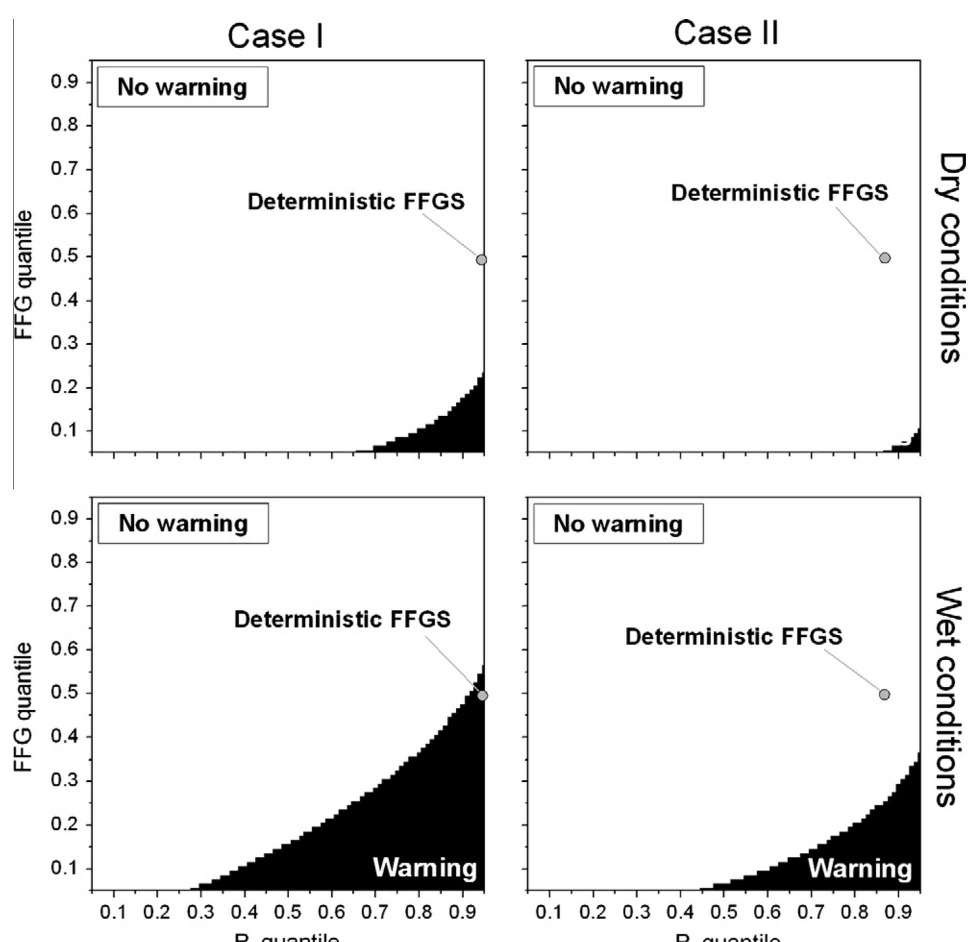

Fig. 5. Plots describing the regions in which warning should or should not be issued according to the different *FFG* and R_r quantiles for both rainfall events and under dry and wet conditions.

accumulation maps averaged over HRAP grid and accumulated them to obtain three-hourly radar-rainfall fields. Table 2 summarizes the values of *FFG* and R_r, averaged over the selected basin, under wet and dry conditions, for the two selected cases.

Results

In the current operational application of the FFGS, *FFG* values are compared with radar-rainfall estimates or rainfall nowcasts (produced either using radar data or independently of radar data through nowcast models), and this information is used with other local information to make decisions concerning whether or not to issue a flash flood watch or warning for a particular basin and a particular time period. For the purposes of this paper we will assume that the forecaster has available only the FFGS information to make a decision. According to the current system, a flash flood warning or watch should (not) be issued for Cases I in wet (dry) conditions, while for Case II *FFG* is always larger than the radar-rainfall value (Table 2). However, under these conditions, the uncertainties associated with radar-rainfall estimation and the hydraulic and hydrologic components of the system are neglected.

Similar to Ntelekos et al. (2006), we evaluate a scenario in which we account for the uncertainties in *FFG* but we still consider the radar-rainfall estimates error-free (probabilistic in *FFG* and deterministic in R_r). As mentioned above, the uncertainties in *FFG* for this basin can be described by a lognormal distribution with mean and variance from Table 11 in Ntelekos et al. (2006). Fig. 3 shows the results from our analysis. Case I resulted in the largest rainfall over the basin. This higher value is highlighted by a higher probability of exceedance of *FFG* by the rainfall. Under dry conditions, the probability that a flash flood would occur is equal to 0.31 and 0.05 for Cases I and II, respectively. This probability increases under wet conditions: it is equal to 0.64 and 0.26, respectively for Cases I and II.

Another possible scenario is that *FFG* is treated as deterministic (error-free), while we account for the uncertainties in radar estimates of rainfall (probabilistic in R_r). In Fig. 4, we show the results for this scenario for the two cases and the different initial soil moisture conditions. Under these conditions, the probability that the true areal ground rainfall is larger than or equal to *FFG* is smaller than 0.01 for both cases under dry conditions, and it increases to 0.09 and 0.01 for Cases I and II under wet conditions. These probabilities are consistently smaller than those obtained under probabilistic *FFG* and deterministic R_r. Ciach et al. (2007) found that the Oklahoma City radar had the tendency to overestimate rainfall (in terms of both unconditional and conditional biases). This aspect could explain the small probabilities of exceedance of *FFG* by the basin-average rainfall (consult also Georgakakos (2005) for additional results regarding the impact of bias in rainfall estimates on the FFGS).

The final step is the development of a scenario in which we account for both radar-rainfall and *FFG* uncertainties. As mentioned above, we no longer compare a single deterministic value with another one or against a distribution, as in the previous three scenarios. We compare a certain quantile q_{FFG} from the distribution of *FFG*

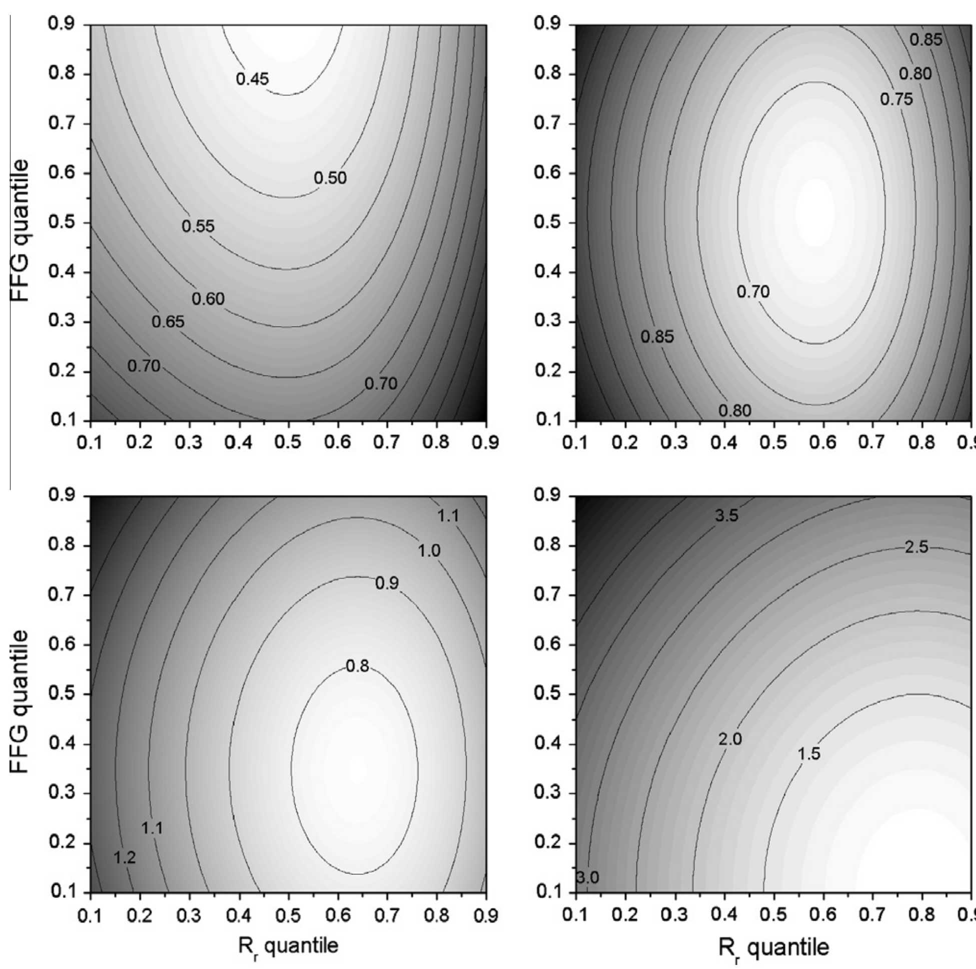

Fig. 6. Plots with the risk function (arbitrary units) for a damage ratio of 0.5 (upper-left panel), 1.0 (upper-right panel), 1.5 (lower-left panel), and 5 (lower-right panel). The damage ratio is defined as the ratio between damage associated with missed warnings and the damage associated with false alarms.

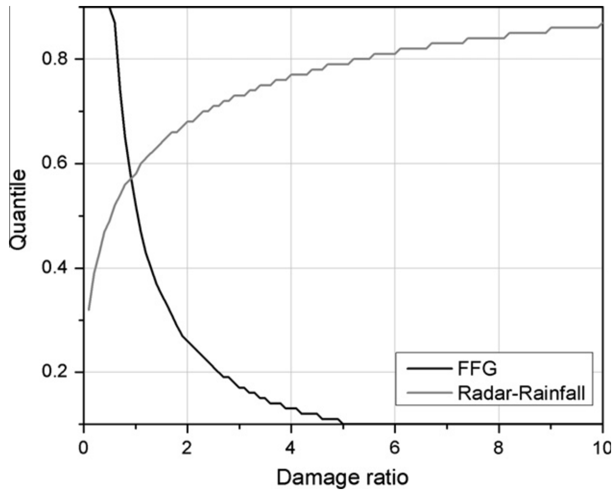

Fig. 7. Plots with the *FFG* (black line) and R_r (grey line) quantiles as a function of the damage ratio. The damage ratio is defined as the ratio between damage associated with missed warnings and the damage associated with false alarms.

with one from the distribution of the radar data q_{RR}. In this way, we can generate plots in which we evaluate whether q_{FFG} is smaller (flash flood warning or watch issued) or larger (flash flood warning or watch not issued) than q_{RR}. In Fig. 5, we present the results for both of the cases under both dry and wet conditions. As expected, as the value of the *FFG* quantile increases, there is less chance that a flash flood warning needs to be issued. The opposite occurs for increasing R_r quantiles.

The next step is the selection of the *FFG* and radar-rainfall quantiles. As mentioned before, we should have information about the probability of missed warnings and false alarms. In this study, we have assumed that it can be described as in Fig. 1. Moreover, we also assume to have information about the "cost" associated with false alarms and missed warnings. Given these two pieces of information we can compute *riskFF* for any *FFG* and radar-rainfall quantiles. For simplicity, we have assumed that the "cost" associated with false alarms is a multiple of the "cost" associated with missed warnings (the "damage ratio" represents the ratio of the missed warning cost to the false alarm cost). In Fig. 6 we have summarized the risk plots for a damage ratio of 0.5 (upper-left panel), 1.0 (upper-right panel), 1.5 (lower-left panel), and 5 (lower-right panel). Given the values in Fig. 1 and for a damage ratio of 1 (same "cost" associated with missed warnings and false alarms), we have that the minimum value of *riskFF* corresponds to a quantile value of 0.58 for radar-rainfall and 0.52 for the *FFG*. For increasing values of the damage ratio, the most "convenient" (see Martina et al. (2006)) combination of quantiles moves to lower *FFG* quantiles and higher radar-rainfall quantiles, reflecting the higher "cost" associated with a missed warning than with a false alarm (Fig. 7). On the other hand, if the "cost" associated with false alarms is larger than the "cost" associated with missed warnings, then the minimum value of *riskFF* moves to lower q_{RR} values and larger q_{FFG} values (Fig. 7).

In this example, the selection of q_{FFG} and q_{RR} is a function of the damage ratio. However, once the "cost" of false alarms and missed warnings is established, single values of these quantiles can be obtained. Once we have q_{FFG} and q_{RR}, we can then use Fig. 5 to help decide about issuing (or not) a flash flood warning or watch.

Discussion and conclusions

The current FFGS used by the NWS to help in issuing flash flood warning or watches is a deterministic system and can be roughly broken down into two main components: the computation of the *FFG* values and the radar-rainfall estimation. However, both components are affected by errors that need to be characterized and accounted for to provide users and decision makers with the tools to make an informed decision. Ntelekos et al. (2006) showed how the uncertainties associated with *FFG* could be described by a lognormal distribution, while Ciach et al. (2007) developed an empirically-based radar-rainfall error model. Using the analytical formulation by Villarini et al. (2009a) to account for radar-rainfall uncertainties, the simplified FFGS was examined under four different scenarios and for two different rainfall events: (i) deterministic in both R_r and *FFG* (current system); (ii) deterministic in R_r and probabilistic in *FFG*; (iii) probabilistic in R_r and deterministic in *FFG*; and (iv) probabilistic in both R_r and *FFG*.

Neglecting the uncertainties and using a deterministic approach would not help in making an optimal decision and would only give an "illusion of certainty" as describe in Krzysztofowicz (2001). On the other hand, a system in which the errors are accounted for would allow users to make more informed decisions. In the proposed modification of the FFGS from a deterministic to a probabilistic system, we can generate a plot in which the uncertainties in radar-rainfall and *FFG* are explicitly accounted for. Thus, users would have to decide whether or not to issue a flash flood warning or watch based on the provided uncertainty information and their knowledge and experience of the system and the study area. Nevertheless, it is clear that the decision concerning the selection of q_{FFG} and q_{RR} is critical: low *FFG* (or high R_r) quantiles could result in a large number of false alarms, with economical and social repercussions, while high *FFG* (or low R_r) quantiles could result in unannounced flash floods, with a potentially significant human toll. From an economic standpoint, the choice could be driven by the use of cost-loss models, which could help optimizing the decision making process (e.g., Mylne, 2002). Therefore, the selection of the quantiles should be based on a detailed cost-benefit analysis resulting from different q_{FFG}–q_{RR} scenarios and should consider the land use and population of the area. In this study, we have provided some indications on how to select the most "convenient" values of quantiles for both *FFG* and radar-rainfall. The proposed results are based on assumptions concerning the probability of missed warnings and the probability of false alarms. To evaluate the performance of the methods advanced here, real data from historical flash floods could be used as a "calibration tool" towards the selection of the quantile to use in the decision process for particular regions and lead times. The development of databases of observed flash flood events in the US is an important undertaking for this purpose (see discussion in Carpenter et al. (2007)).

To complicate the matter, there is also interdependency between the two components, since initial soil moisture conditions are computed using the radar-rainfall estimates. Future studies should investigate the impact of radar-rainfall uncertainties on the initial soil moisture conditions, and how these uncertainties propagate through the FFGS. Initial steps illustrating methodological aspects along this direction have been made in Georgakakos (2005) but significant research remains to be done for the development of operationally useful flash flood warning systems.

Admittedly, our results are based on only two cases, assuming that the parameters of the error model estimated for the Oklahoma City radar are valid for the Tulsa radar as well. Nevertheless, this is an important advancement towards the transition from a deterministic to probabilistic flash flood alert system. More effort is required to extend this formulation to areas characterized by different rainfall regimes and/or more complex orography: the characterization of the errors in *FFG* and radar-rainfall for different areas or radars is not yet available and should be the object of future studies. In particular, to characterize the uncertainties in *FFG* for a given catchment, we would need information derived from GIS (drainage area, main stream length, Horton's length ratio)

and from regional relationships (channel top width and hydraulic depth at the outlet, and local channel slope), as in Carpenter et al. (1999). For describing the uncertainties in radar-rainfall according to the model by Ciach et al. (2007), we would need a large dataset of rain gage and radar-rainfall data to compute the parameters of the systematic and random components of the error model. More generally, the broad development of validation programs for flash-flood forecasting could serve the dual roles of providing quantitative assessments of forecast accuracy and data sets that would be useful for enhancing models of *FFG* error structure.

Error characteristics for urban catchments can exhibit properties that are especially difficult to characterize. Javier et al. (2007) illustrated the role of urban infrastructure in determining error properties of flash-flood forecasting components. Advances in monitoring land surface properties (see, for example, Horritt and Bates, 2002; Lane et al., 2003) for urban catchments and integration of new data sources into representations of model error structure are important research areas.

Finally, given the dependence of the radar-rainfall error model components on radar-rainfall values, we can characterize the uncertainties associated with extreme precipitation events, which are generally the most interesting in these applications. However, modeling of radar-rainfall uncertainties for this type of events may require further efforts and should be investigated in separate studies.

Acknowledgments

The first author was supported by NASA Headquarters under the Earth Science Fellowship Grant NNX06AF23H while a graduate student at the University of Iowa. The authors wish to thank two anonymous reviewers and the associate editor Dr. Marco Borga for their helpful comments.

References

Allen, P.B., Naney, J.W., 1991. Hydrology of the Little Washita River Watershed, Oklahoma: Data and Analysis. US Department of Agriculture, Agricultures Research Service.

AMS, 2000. Policy statement: prediction and mitigation of flash floods. Bulletin of the American Meteorological Society 81 (6), 1338–1340.

Ashley, S.T., Ashley, W.S., 2008. Flood fatalities in the United States. Journal of Applied Meteorology and Climatology 47 (3), 805–818.

Baeck, M.L., Smith, J.A., 1998. Rainfall estimation by the WSR-88D for heavy rainfall events. Weather and Forecasting 13 (2), 416–436.

Borga, M., Boscolo, P., Zanon, F., Sangati, M., 2007. Hydrometeorological analysis of the 29 August 2003 flash flood in the Eastern Italian Alps. Journal of Hydrometeorology 8 (5), 1049–1067.

Brock, F.V. et al., 1995. The Oklahoma mesonet: a technical overview. Journal of Atmospheric and Oceanic Technology 12 (1), 5–19.

Burnash, R.J.C., 1995. Computer Models of Watershed Hydrology. Water Resources Publication.

Carpenter, T.M., Georgakakos, K.P., 1993. GIS-based Procedure in Support of Flash Flood Guidance. The University of Iowa, Iowa City, Iowa.

Carpenter, T.M., Sperfslage, J.A., Georgakakos, K.P., Sweeney, T., Fread, D.L., 1999. National threshold runoff estimation utilizing GIS in support of operational flash flood warning systems. Journal of Hydrology 224 (1–2), 21–44.

Carpenter, T.M. et al., 2007. Surveying flash flood response in mountainous streams. EOS Transactions American Geophysical Union 88 (6).

Chancibault, K., Anquetin, S., Ducrocq, V., Saulnier, G.M., 2006. Hydrological evaluation of high-resolution precipitation forecasts of the Gard flash-flood event (8–9 September 2002). Quarterly Journal of the Royal Meteorological Society 132 (617), 1091–1117.

Chang, C.B., 1998. A case study of excessive rainfall forecasting. Meteorology and Atmospheric Physics 66 (3–4), 215–227.

Choi, H., Kang, I.J., Chung, Y.H., Yoon, H.J., 2007. Uncertainty analysis of flash-flood system using remote sensing and a geographic information system based on GcIUH in the Yeongdeok Basin, Korea. International Journal of Remote Sensing 28 (24), 5551–5565.

Ciach, G.J., Krajewski, W.F., Villarini, G., 2007. Product-error-driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data. Journal of Hydrometeorology 8 (6), 1325–1347.

Collier, C.G., 2007. Flash flood forecasting: what are the limits of predictability? Quarterly Journal of the Royal Meteorological Society 133 (622), 3–23.

Crum, T.D., Alberty, R.L., 1993. The WSR-88D and the WSR-88D operational support facility. Bulletin of the American Meteorological Society 74 (9), 1669–1687.

Delrieu, G. et al., 2005. The catastrophic flash-flood event of 8–9 September 2002 in the Gard region, France. A first case study for the Cevennes–Vivarais mediterranean hydrometeorological observatory. Journal of Hydrometeorology 6 (1), 34–52.

Downton, M.W., Barnard Miller, J.Z., Pielke, R.A., 2005. Reanalysis of US national weather service flood loss database. Natural Hazards Review 6 (1), 13–22.

Ferraris, L., Rudari, R., Siccardi, F., 2002. The uncertainty in the prediction of flash floods in the northern Mediterranean environment. Journal of Hydrometeorology 3 (6), 714–727.

Fulton, R.A., Breidenbach, J.P., Seo, D.J., Miller, D.A., O'Bannon, T., 1998. The WSR-88D rainfall algorithm. Weather and Forecasting 13 (2), 377–395.

Georgakakos, K.P., 2005. Modern operational flash flood warning systems based on flash flood guidance theory: performance evaluation. In: Proceedings of the International Conference on Innovation, Advances and Implementation of Flood Forecasting Technology. Bergen-Tromso, Norway. pp. 1–10.

Georgakakos, K.P., 2006. Analytical results for operational flash flood guidance. Journal of Hydrology 317 (1–2), 81–103.

Germann, U., Galli, G., Boscacci, M., Bolliger, M., 2006. Radar precipitation measurement in a mountainous region. Quarterly Journal of the Royal Meteorological Society 132 (618), 1669–1692.

Horritt, M.S., Bates, P.D., 2002. Evaluation of 1D and 2D numerical models for predicting river flood inundation. Journal of Hydrology 268 (1–4), 87–99.

Javier, J.R.N., Smith, J.A., Meierdiercks, K.L., Baeck, M.L., Miller, A.J., 2007. Flash flood forecasting for small urban watersheds in the Baltimore metropolitan region. Weather and Forecasting 22 (6), 1331–1344.

Jessup, S.M., DeGaetano, A.T., 2008. A statistical comparison of the properties of flash flooding and nonflooding precipitation events in portions of New York and Pennsylvania. Weather and Forecasting 23 (1), 114–130.

Klazura, G.E., Imy, D.A., 1993. A description of the initial set of analysis products available from the NEXRAD WSR-88D system. Bulletin of the American Meteorological Society 74 (7), 1293–1311.

Krajewski, W.F., 2008. Hydro-NEXRAD: an updated overview and metadata analysis. World Environmental and Water Resources Congress 2008. Honolulu, Hawaii.

Krajewski, W.F. et al., 2007a. Towards Better Utilization of NEXRAD Data in Hydrology: An Overview of Hydro-NEXRAD. World Environmental and Water Resources Congress 2007, Tampa, Florida.

Krajewski, W.F. et al., 2007b. Hydro-NEXRAD Radar-Rainfall Estimation Algorithm Development, Testing, and Evaluation. World Environmental and Water Resources Congress 2007, Tampa, Florida.

Krajewski, W.F., Smith, J.A., 2002. Radar hydrology: rainfall estimation. Advances in Water Resources 25 (8–12), 1387–1394.

Krzysztofowicz, R., 2001. The case for probabilistic forecasting in hydrology. Journal of Hydrology 249 (1–4), 2–9.

Lane, S.N., James, T.D., Pritchard, H., Saunders, M., 2003. Photogrammetric and laser altimetric reconstruction of water levels for extreme flood event analysis. Photogrammetric Record 18 (104), 293–307.

Martina, M.L.V., Todini, E., Libralon, A., 2006. A Bayesian decision approach to rainfall thresholds based flood warning. Hydrology and Earth System Sciences 10 (3), 413–426.

Morin, E., Jacoby, Y., Navon, S., Bet-Halachmi, E., 2009. Towards flash-flood prediction in the dry Dead Sea region utilizing radar rainfall information. Advances in Water Resources 32, 1066–1076.

Mylne, K.R., 2002. Decision-making from probability forecasts based on forecast value. Meteorological Applications 9 (3), 307–315.

Norbiato, D., Borga, M., Dinale, R., 2009. Flash flood warning in ungauged basins by use of the flash flood guidance and model-based runoff thresholds. Meteorological Applications 16 (1), 65–75.

Norbiato, D., Borga, M., Esposti, S.D., Gaume, E., Anquetin, S., 2008. Flash flood warning based on rainfall thresholds and soil moisture conditions: an assessment for gauged and ungauged basins. Journal of Hydrology 362 (3–4), 274–290.

NRC, 2005. Flash Flood Forecasting Over Complex Terrain, Washington, DC.

Ntelekos, A.A., Georgakakos, K.P., Krajewski, W.F., 2006. On the uncertainties of flash flood guidance. Toward probabilistic forecasting of flash floods. Journal of Hydrometeorology 7 (5), 896–915.

O'Connor, J.E., Costa, J.E., 2004. Spatial distribution of the largest rainfall–runoff floods from basins between 2.6 and 26,000 km(2) in the United States and Puerto Rico. Water Resources Research 40 (1).

Ogden, F.L. et al., 2000. Hydrologic analysis of the Fort Collins, Colorado, flash flood of 1997. Journal of Hydrology 228 (1–2), 82–100.

Petersen, W.A. et al., 1999. Mesoscale and radar observations of the Fort Collins flash flood of 28 July 1997. Bulletin of the American Meteorological Society 80 (2), 191–216.

Pielke, R.A., Downton, M.W., 2000. Precipitation and damaging floods: trends in the United States, 1932–97. Journal of Climate 13 (20), 3625–3637.

Pontrelli, M.D., Bryan, G., Fritsch, J.M., 1999. The Madison County, Virginia, flash flood of 27 June 1995. Weather and Forecasting 14 (3), 384–404.

Reed, S.M., Maidment, D.R., 1999. Coordinate transformation for using NEXRAD data in GIS-based hydrologic modeling. Journal of Hydraulic Engineering 4, 174–182.

Reggiani, P., Weerts, A.H., 2008. Probabilistic quantitative precipitation forecast for flood prediction: an application. Journal of Hydrometeorology 9 (1), 76–95.

Rodriguez-Iturbe, I., Gonzalezsanabria, M., Bras, R.L., 1982a. A geomorphoclimatic theory of the instantaneous unit-hydrograph. Water Resources Research 18 (4), 877–886.
Rodriguez-Iturbe, I., Sanabria, M.G., Caamano, G., 1982b. On the climatic dependence of the IUH – a rainfall–runoff analysis of the Nash model and the geomorphoclimatic theory. Water Resources Research 18 (4), 887–903.
Rodriguez-Iturbe, I., Valdes, J.B., 1979. Geomorphologic structure of hydrologic response. Water Resources Research 15 (6), 1409–1420.
Scofield, R.A., Kuligowski, R.J., 2003. Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. Weather and Forecasting 18 (6), 1037–1051.
Senesi, S., Bougeault, P., Cheze, J.L., Cosentino, P., Thepenier, R.M., 1996. The Vaison-La-Romaine flash flood: mesoscale analysis and predictability issues. Weather and Forecasting 11 (4), 417–442.
Siccardi, F., Boni, G., Ferraris, L., Rudari, R., 2005. A hydrometeorological approach for probabilistic flood forecast. Journal of Geophysical Research – Atmospheres 110 (D5).
Smith, J.A., Baeck, M.L., Morrison, J.E., Sturdevant-Rees, P., 2000. Catastrophic rainfall and flooding in Texas. Journal of Hydrometeorology 1 (1), 5–25.
Smith, J.A., Baeck, M.L., Zhang, Y., Doswell, C.A., 2001. Extreme rainfall and flooding from supercell thunderstorms. Journal of Hydrometeorology 2 (5), 469–489.
Taramasso, A.C., Gabellani, S., Parodi, A., 2005. An operational flash-flood forecasting chain applied to the test cases of the EU project HYDROPTIMET. Natural Hazards and Earth System Sciences 5 (5), 703–710.
Villarini, G., Krajewski, W.F., 2009a. Empirically based modelling of radar-rainfall uncertainties for a C-band radar at different time-scales. Quarterly Journal of the Royal Meteorological Society 135 (643), 1424–1438.
Villarini, G., Krajewski, W.F., 2009b. Inference of spatial scaling properties of rainfall: impact of radar rainfall estimation uncertainties. IEEE Geoscience and Remote Sensing Letters 6 (4), 812–815.
Villarini, G., Krajewski, W.F., 2010a. Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. Surveys in Geophysics 31 (1), 107–129.
Villarini, G., Krajewski, W.F., 2010b. Sensitivity studies of the models of radar-rainfall uncertainties. Journal of Applied Meteorology and Climatology 49 (2), 288–309.
Villarini, G., Krajewski, W.F., Ciach, G.J., Zimmerman, D.L., 2009a. Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates. Water Resources Research 45, W01404. doi:10.1029/2008WR006946.
Villarini, G., Krajewski, W.F., Smith, J.A., 2009b. New paradigm for statistical validation of satellite precipitation estimates: application to a large sample of the TMPA 0.25 degrees 3-hourly estimates over Oklahoma. Journal of Geophysical Research – Atmospheres 114, D12106. doi:10.1029/2008JD011475.
Villarini, G., Serinaldi, F., Krajewski, W.F., 2008. Modeling radar-rainfall estimation uncertainties using parametric and non-parametric approaches. Advances in Water Resources 31 (12), 1674–1686.
Weaver, J.F., Gruntfest, E., Levy, G.M., 2000. Two floods in Fort Collins, Colorado: learning from a natural disaster. Bulletin of the American Meteorological Society 81 (10), 2359–2366.
Wilson, J.W., Brandes, E.A., 1979. Radar measurement of rainfall – a summary. Bulletin of the American Meteorological Society 60 (9), 1048–1058.
Zehe, E., Bloschl, G., 2004. Predictability of hydrologic response at the plot and catchment scales: role of initial conditions. Water Resources Research 40 (10).
Zehe, E., Elsenbeer, H., Lindenmaier, F., Schulz, K., Bloschl, G., 2007. Patterns of predictability in hydrological threshold systems. Water Resources Research 43 (7), 12.
Zehe, E., Sivapalan, M., 2009. Threshold behaviour in hydrological systems as (human) geo-ecosystems: manifestations, controls, implications. Hydrology and Earth System Sciences 13 (7), 1273–1297.